

A novel technique to incorporate structural prior information into multi-modal tomographic reconstruction

Daniil Kazantsev^{1,2}, Sébastien Ourselin³, Brian F Hutton⁴, Katherine J Dobson^{1,2}, Anders P Kaestner⁵, William R B Lionheart⁶, Philip J Withers^{1,2}, Peter D Lee^{1,2} and Simon R Arridge³

¹ The Manchester X-Ray Imaging Facility, School of Materials, The University of Manchester, Manchester, M13 9PL, UK

² The Manchester X-Ray Imaging Facility, Research Complex at Harwell, Didcot, Oxfordshire, OX11 0FA, UK

³ Centre for Medical Image Computing, University College London, London NW1 9EE, UK

⁴ Institute of Nuclear Medicine, UCL and University College London Hospitals NHS Trust, 235 Euston Road (T-5), London NW1 2BU, UK

⁵ Spallation Neutron Source Division, Paul Scherrer Institut (PSI), 5232, Villigen, Switzerland

⁶ School of Mathematics, Alan Turing Building, The University of Manchester, M13 9PL, UK

E-mail: daniil.kazantsev@manchester.ac.uk

Received 22 November 2013, revised 16 February 2014

Accepted for publication 18 February 2014

Published 19 May 2014

Abstract

There has been a rapid expansion of multi-modal imaging techniques in tomography. In biomedical imaging, patients are now regularly imaged using both single photon emission computed tomography (SPECT) and x-ray computed tomography (CT), or using both positron emission tomography and magnetic resonance imaging (MRI). In non-destructive testing of materials both neutron CT (NCT) and x-ray CT are widely applied to investigate the inner structure of material or track the dynamics of physical processes. The potential benefits from combining modalities has led to increased interest in iterative reconstruction algorithms that can utilize the data from more than one imaging mode simultaneously. We present a new regularization term in iterative reconstruction that enables information from one imaging modality to be used as a structural prior to improve resolution of the second modality. The regularization term is based on a modified anisotropic tensor diffusion filter, that has shape-adapted smoothing properties. By considering the underlying



orientations of normal and tangential vector fields for two co-registered images, the diffusion flux is rotated and scaled adaptively to image features. The images can have different greyscale values and different spatial resolutions. The proposed approach is particularly good at isolating oriented features in images which are important for medical and materials science applications. By enhancing the edges it enables both easy identification and volume fraction measurements aiding segmentation algorithms used for quantification. The approach is tested on a standard denoising and deblurring image recovery problem, and then applied to 2D and 3D reconstruction problems; thereby highlighting the capabilities of the algorithm. Using synthetic data from SPECT co-registered with MRI, and real NCT data co-registered with x-ray CT, we show how the method can be used across a range of imaging modalities.

Keywords: iterative tomographic reconstruction, hybrid modalities, emission and neutron tomography, regularization methods, anisotropic diffusion

(Some figures may appear in colour only in the online journal)

1. Introduction

Biomedical imaging modalities, such as single photon emission computed tomography (SPECT) and positron emission tomography (PET) provide information about the distribution of radioisotopes and tracers in living tissue. These distributions allow physicians to determine certain chemical properties of the tissue [1]. The major drawback of the functional modalities is low signal-to-noise ratio and low spatial resolution of the reconstructed images. On the other hand, modalities such as x-ray CT and magnetic resonance imaging (MRI) are able to achieve a superior spatial resolution of anatomical structure.

The availability of additional, independent information from hybrid multi-modal scanners (SPECT/x-ray CT, PET/MRI, PET/x-ray CT etc) can improve resolution of one image by referring to the prominent edges in the other [2, 3]. In materials science, the same sample volume can be scanned using multiple imaging systems, such as x-ray CT and neutron CT (NCT).

For reconstruction of SPECT and PET data iterative statistical methods such as the maximum-likelihood expectation maximization (MLEM) algorithm [4] are commonly used, since the reconstruction problem is ill-posed and ill-conditioned and reconstruction requires additional regularization to ensure well-posedness. From the Bayesian perspective, imposing desirable properties (e.g. smoothness) onto a solution leads to the characterization of an *a priori* probability distribution [5]. Iteration then uses the maximum *a posteriori* probability estimates or penalized likelihood methods to find a trade-off between the data term and the noise suppressing term.

In this paper we propose a new method to incorporate multi-modal information into the smoothing process, and apply this method to very two different problems: the determination of radiopharmaceutical distribution in human brain, and for locating liquids in gravel beds. Our proposed method is based on recently developed diffusion filters for tomographic regularization [6, 7]. By modifying the conventional diffusivity filter [8, 9] we present a novel method of embedding additional prior information into the filtering process. Successful image noise reduction while leaving important features intact is of paramount importance for accurate interpretation of recovered images. Penalty functionals based on first derivatives of images (i.e. gradients) [10, 11], such as total variation [12], lead, under variation, to an Euler–Lagrange equation that is of anisotropic diffusion (AD) [8, 9] type.

Here we use a modified anisotropic diffusivity tensor regularization term which depends on supplementary information. The structural information is collected from a higher resolution dataset of the same volume, and the tensor interpolation performed adaptively to the underlying image features. We present a treatment based on two sources of data, but the method can be generalized to higher order data spaces. To reach the solution a nested two-step algorithm [13] is utilized for iterative optimization of the cost function. The main cost function is decomposed into two sub-problems with a separate optimization of Poisson data fidelity and variational penalty terms. The resulting algorithm consists of a classical MLEM step followed by a second step for the variational problem. This second step is the focus of this contribution.

The approach presented is tested on an image denoising and deblurring problem and then applied to synthetic SPECT/MR data and to material imaging. For the sake of simplicity we consider the 2D case, however the proposed idea can be easily generalized to the 3D reconstruction problem. We introduce hot and cold lesions into the synthetic emission phantom which are absent from MR image. The main goal is to enhance the spatial resolution of emission images without loss of important features, such as introduced lesions. A real data experiment is presented where noisy neutron data with a limited number of projections are reconstructed with the aid of supplementary x-ray CT data.

The paper is organized as follows. In section 2, different diffusion strategies are described using the general framework of tensor based AD filtering. In section 3 the image recovery example is used to demonstrate the behaviour of the proposed method. The different problems relating to the choice of parameters are considered. The numerical results are presented for simulated emission tomographic reconstruction (SPECT) and real data arising from a neutron tomography experiment is reconstructed. Section 4 gives a discussion and the conclusions are summarized in section 5.

2. Method

2.1. Functional minimization for nonlinear PDEs

The use of PDEs for image processing tasks, such as restoration, segmentation, deblurring, has been an active area of research since late 1980s [8]. The main feature of these methods is the time evolving nature of the process which decomposes a signal into the set of more global and simpler representations of it [9]. This is regarded as a *scale-space* decomposition which effectively reduces dimensionality of noisy signal by projecting it onto the lower dimension spaces. The main challenge remains suppressing noise while preserving discontinuities related to the valuable features. In terms of evolution equations one needs to find a suitable steady state of the process when the scale-space decomposition represents the most valuable information in the signal.

Many evolution equations can be considered as gradient descent methods for minimizing a suitable energy functional [9]. To restore the image u one needs to minimize:

$$\min_u E(u) = \int_{\Omega} \phi(|\nabla u|) \, d\Omega, \quad (1)$$

where $\phi(|\nabla u|)$ is an increasing function of the gradients magnitude $|\nabla u| = \sqrt{u_x^2 + u_y^2}$ with $u_x = \frac{\partial u}{\partial x}$ and $u_y = \frac{\partial u}{\partial y}$ respectively.

The Euler–Lagrange equations of $E(u)$ are:

$$\frac{\partial E(u)}{\partial u} = -\nabla \cdot \left(\frac{\phi'(|\nabla u|)}{|\nabla u|} \nabla u \right), \quad (2)$$

where $\nabla \cdot (\mathbf{v}) = v_x + v_y$ is the divergence operator. For functional minimization one can perform gradient descent as a continuous time evolution:

$$\partial_t u = -\frac{\partial E(u)}{\partial u} = \nabla \cdot \left(\frac{\phi'(|\nabla u|)}{|\nabla u|} \nabla u \right), \quad u_{t=0} = \hat{u}(\mathbf{x}), \quad (3)$$

where initialization $\hat{u}(\mathbf{x})$ is a noisy image. Therefore one can consider the diffusion process (3) to be equivalent to minimization of the functional (1).

Equation (3) is a Perona–Malik diffusion evolution [8], where diffusion coefficient is defined as:

$$\varphi(|\nabla u|) = \frac{\phi'(|\nabla u|)}{|\nabla u|}. \quad (4)$$

The choice of function $\phi(|\nabla u|)$ can lead to different linear and nonlinear models [9]. In this paper we use the following convex functional: $\phi(s) = -\frac{\epsilon^2}{2} \exp(-s^2/\epsilon^2)$; $\phi'(s) \geq 0$, $\phi''(s) \geq 0$, which leads to the edge preserving function [10]:

$$\varphi(s) = \exp(-s^2/\epsilon^2), \quad (5)$$

here threshold parameter ϵ controls the strength of diffusion.

One can also re-write equation (3) with a nonlinear isotropic tensor \mathbf{D} :

$$\partial_t u = \nabla \cdot (\mathbf{D} \nabla u), \quad u_{t=0} = \hat{u}(\mathbf{x}), \quad (6)$$

where $\mathbf{D} = \varphi(|\nabla u|)\mathbf{I}$, and \mathbf{I} is the identity matrix.

2.2. Tensor based anisotropic diffusion filtering

In [9], Weickert replaced the isotropic diffusivity tensor in equation (6) with anisotropic diffusivity tensor $\mathbf{D}(\mathbf{J}_\rho)$ which is based on the structure tensor \mathbf{J}_ρ . Anisotropic behaviour of the diffusivity tensor $\mathbf{D}(\mathbf{J}_\rho)$ can enhance geometrically valuable features of an image by differently penalizing normal and tangential directions of smoothing.

The structure tensor can be constructed as a symmetric, positive semi-definite matrix:

$$\mathbf{J}_\rho(\nabla u_\sigma)(\mathbf{x}) = k_\rho * (\nabla u_\sigma \nabla u_\sigma^T) = \begin{bmatrix} k_\rho * u_x^2 & k_\rho * u_x u_y \\ k_\rho * u_x u_y & k_\rho * u_y^2 \end{bmatrix}, \quad (7)$$

where for regularization of the PDE and stable evolution, the image $u(\mathbf{x})$ is convolved with a Gaussian kernel k_σ , σ is a *differentiation scale*: $u_\sigma(\mathbf{x}) = (k_\sigma * u)(\mathbf{x})$. The local information is averaged by convolving component-wise $\nabla u_\sigma \nabla u_\sigma^T$ with a Gaussian kernel k_ρ , where ρ is an *integration scale* which controls the size of the neighbourhood $O(\rho)$ with dominant orientations of the image gradient [10].

Principal axis transformation of (7) gives the orthonormal eigenvectors $\mathbf{v}_1, \mathbf{v}_2$, with \mathbf{v}_1 parallel to

$$\left(j_{22} - j_{11} + \sqrt{(j_{22} - j_{11})^2 + 4j_{12}^2} \right),$$

where j_{lk} are elements of the matrix $\mathbf{J}_\rho(\nabla u_\sigma)$. The corresponding eigenvalues are given as:

$$\eta_1 = \frac{1}{2}(j_{11} + j_{22} + \sqrt{(j_{11} - j_{22})^2 + 4j_{12}^2}); \quad \eta_2 = \frac{1}{2}(j_{11} + j_{22} - \sqrt{(j_{11} - j_{22})^2 + 4j_{12}^2}).$$

The eigenvalues (averaged by the scale parameter ρ) convey the level of contrast along the given eigenvectors within the neighbourhood of size $O(\rho)$. Note also that:

$$\lim_{\rho \rightarrow 0} \mathbf{v}_1 = \mathbf{n} = \frac{\nabla u_\sigma}{|\nabla u_\sigma|}; \quad \lim_{\rho \rightarrow 0} \mathbf{v}_2 = \boldsymbol{\tau} = \frac{\nabla u_\sigma^\perp}{|\nabla u_\sigma|} \quad (8)$$

where \mathbf{n} and $\boldsymbol{\tau}$ are unit vectors, respectively, parallel and orthogonal to the gradient. The vector \mathbf{v}_1 indicates the orientation maximizing the gray-value fluctuations, while \mathbf{v}_2 gives the preferred direction of smoothing (along the level sets). The eigenvalues are descriptors of the geometrical information of \mathbf{x} , for almost constant regions $\eta_1 \cong \eta_2 \cong 0$ and $\eta_1 \gg \eta_2$ along image edges.

Having defined the structure tensor (7), the diffusion tensor (DT) is given as:

$$\mathbf{D}(\mathbf{J}_\rho) = [\mathbf{v}_1 \ \mathbf{v}_2] \begin{bmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \end{bmatrix} = \begin{bmatrix} D_{11} & D_{12} \\ D_{12} & D_{22} \end{bmatrix}, \quad (9)$$

$$D_{ij} = \sum_{n=1,2} \gamma_n v_{ni} v_{nj}. \quad (10)$$

The DT (9) has two new eigenvalues $\gamma_{1,2}$, which define the strength of smoothing in the preferred directions $\mathbf{v}_{1,2}$. The different definitions for γ_n relating to the edge enhancing diffusion (EED) and the coherence enhancing diffusion were proposed by Weickert [14]. Here we use the EED approach with the edge preserving function (5), then the eigenvalues $\gamma_{1,2}$ related to normal and tangential directions respectively are given as:

$$\begin{cases} \gamma_1 = \varphi(|\nabla u|) = \exp(-|\nabla u|^2/\epsilon^2), \\ \gamma_2 = 1. \end{cases} \quad (11)$$

The threshold parameter ϵ controls the strength of diffusion in normal direction. One would like to encourage smoothing in the tangential direction (the direction of level sets), therefore $\gamma_2 = 1$. The DT \mathbf{D} rotates and scales the flux in order to adapt to the underlying geometrical configurations of image $u(\mathbf{x})$.

The time evolving tensor based AD equation with boundary conditions is given as:

$$\begin{cases} \partial_t u = \nabla \cdot (\mathbf{D}(\mathbf{J}_\rho(\nabla u_\sigma)) \nabla u_\sigma) & \text{in } \Omega \times (0, T), \\ \partial_N u = 0 & \text{on } \partial\Omega \times (0, T), \\ u(\mathbf{x}, t = 0) = \hat{u} & \text{in } \Omega, \end{cases} \quad (12)$$

where N is the unit normal to the boundary $\partial\Omega$.

2.3. Incorporating additional information into the diffusivity tensor

In embedding information from image $\mu(\mathbf{x})$ into image $\lambda(\mathbf{x})$ by means of the tensor based diffusion filtering (12), the goal is to enhance λ (which is assumed to be noisy and blurry) using the supplementary information in μ , while preserving the prominent features in λ . The images can have different greyscale values and different spatial resolutions. In this case the spatial resolution of μ is assumed to be higher than λ , although this is not necessary (see section 4).

If we assume that the DTs \mathbf{D}_μ and \mathbf{D}_λ are defined for images μ and λ respectively using equation (9), then following the framework for DT calculus [15] we can use tensor interpolation to construct a *combined* DT (CDT) $\mathbf{D}_{\lambda,\mu}$.

The standard approach is to use the geometrical and intensity information from both images while constructing $\mathbf{D}_{\lambda,\mu}$ is as an arithmetic interpolation:

$$\mathbf{D}_{\lambda,\mu} = s\mathbf{D}_\lambda + (1-s)\mathbf{D}_\mu; \quad s(\mathbf{x}) \in [0, 1], \quad (13)$$

where $s(\mathbf{x})$ is a spatially variant parameter that controls the trade-off between the two tensors. It is important to note that this choice is purely empirical and other techniques for tensor fusing can be used instead [15, 16].

To emphasize the importance of the parameter s in the interpolation procedure (13), we first consider the desirable properties of this parameter:

Property 1. If $|\nabla\mu| \rightarrow 0$, then $s \rightarrow 1$, else if $|\nabla\lambda| \rightarrow 0$, then $s \rightarrow 0$.

Property 2. If $|\nabla\mu|$ and $|\nabla\lambda| \rightarrow 0$, then $\mathbf{D}_{\lambda,\mu} \rightarrow \mathbf{I}$.

Properties 1 and 2 deal with uniform areas on images and special cases have to be considered for the stable performance of the CDT method. Property 1 focuses on the fact that when local information is absent in one image it might exist in the other image. Property 2 states that when the intensity level is uniform for both images the tensor becomes isotropic.

One should be careful when imposing a dependence on gradient magnitudes alone [17], while smoothing it can introduce undesirable edges which are not common to both images. One possible solution will be to evaluate a gradient orientation angle for every \mathbf{x} in image u as:

$$\theta(u(\mathbf{x})) = \arctan\left(\frac{u_y}{u_x}\right) + \frac{\pi}{2}, \quad (14)$$

however this method is very sensitive to noise [18]. In this paper we use a more robust way by considering orientations provided by the structure tensor $\mathbf{J}_\rho(\nabla u_\sigma)$ (7). Using the eigendirections from μ and λ the angle $\hat{\theta}$ between two vector bases can be estimated as:

$$\hat{\theta} = \arccos(\mathbf{v}_1(\mu) \cdot \mathbf{v}_1(\lambda)). \quad (15)$$

We would like to penalize smaller angles of $\hat{\theta}$ because in this case the vectors become parallel and edges tend to have the same orientation. For estimation of $s(\mathbf{x})$ we use a two-step procedure:

Step 1. Find the common edges of μ and λ by thresholding of the edge preserving functions $\varphi_{\epsilon_\mu}(|\nabla\mu|)$, $\varphi_{\epsilon_\lambda}(|\nabla\lambda|)$. The edge is said to be common if $\varphi_{\epsilon_\mu}(|\nabla\mu|)$ and $\varphi_{\epsilon_\lambda}(|\nabla\lambda|)$ are less than a small constant ς .

Step 2. Calculate s for the common edges based on the response of $\hat{\theta}$ (15) as:

$$s(\mathbf{x}) = \sin\left(\text{abs}\left(\hat{\theta} - \frac{\pi}{2}\right)\right), \quad (16)$$

since $\hat{\theta} \in (0, \pi)$ we recalculate it to be in $(-\frac{\pi}{2}, \frac{\pi}{2})$ and take the absolute value to achieve the range $(0, \frac{\pi}{2})$. As we wish to penalize angles which are closer to zero we use a sine function. We do not claim this strategy to be optimal, but it is sufficient to demonstrate the applicability of the proposed algorithm (see section 4).

The challenge in incorporating the structural prior is to avoid bias in the λ restoration process, while imposing constraints provided by μ . A similar trade-off problem arises in multi-modal tomographic reconstruction when the image of one tomographic modality is reconstructed using data from the other modality as *a priori* information [6].

2.4. Discretization of the proposed model

The divergence term (12) with the combined tensor (13) are given as:

$$\nabla \cdot (\mathbf{D}_{\lambda,\mu} \nabla \lambda) = \nabla \cdot \begin{pmatrix} D_{11}\partial_x\lambda + D_{12}\partial_y\lambda \\ D_{12}\partial_x\lambda + D_{22}\partial_y\lambda \end{pmatrix}, \quad (17)$$

then:

$$\partial_t \lambda = \partial_x(D_{11}\partial_x\lambda) + \partial_x(D_{12}\partial_y\lambda) + \partial_y(D_{12}\partial_x\lambda) + \partial_y(D_{22}\partial_y\lambda). \quad (18)$$

We use the following asymmetric (left and right derivative) schemes:

$$\begin{aligned} \partial_x(D_{11}\partial_x\lambda) &= \frac{1}{2}((D_{11}[i, j] + D_{11}[i+1, j])(\lambda[i+1, j] - \lambda[i, j]) \\ &\quad - (D_{11}[i-1, j] + D_{11}[i, j])(\lambda[i, j] - \lambda[i-1, j])) + O(h^2) \end{aligned}$$

and use the central differences for the cross terms:

$$\begin{aligned} \partial_y(D_{12}\partial_x\lambda) &\approx \frac{1}{4}((D_{12}[i, j+1](\lambda[i+1, j+1] - \lambda[i-1, j+1]) \\ &\quad - D_{12}[i, j-1](\lambda[i+1, j-1] - \lambda[i-1, j-1])) + O(h^2), \end{aligned}$$

and forward time derivative:

$$\partial_t \lambda = \frac{\lambda(x, y, t + \Delta t) - \lambda(x, y, \Delta t)}{\Delta t}.$$

Here Δt denotes the time discretization constant. By using the approximations above one can use the following explicit numerical scheme:

$$\lambda^{l+1} = \lambda^l + \Delta t (\partial_x (D_{11} \partial_x \lambda^l) + \partial_x (D_{12} \partial_y \lambda^l) + \partial_y (D_{12} \partial_x \lambda^l) + \partial_y (D_{22} \partial_y \lambda^l)), \quad (19)$$

Δt is normally chosen to be a small value (e.g. for linear heat flow $\Delta t \leq 0.25$) to ensure stability of the fully explicit numerical scheme [9]. The iterative process (19) is a time evolving scale space decomposition and the number of iterations l controls the smoothing strength.

2.5. Iterative emission tomography reconstruction using CDT filtering

The image reconstruction task in ET is to recover the unknown radiotracer distribution $\lambda \in \mathbb{R}^N$ which is an N -dimensional vector, while using the measured projection data (sinogram) $g \in \mathbb{R}^M$. In the ET, g follows a Poisson distribution with mean:

$$\bar{g} = P\lambda + s \quad (20)$$

where the projection or system matrix $P : \mathbb{R}^N \rightarrow \mathbb{R}^M$ depends on the system design and the detector array geometry. In this work we do not account for scatter effects s , but geometrical characteristics, such as intrinsic resolution of the detection system and collimator resolution, were included in P .

To reconstruct the image λ from the measured data g , the following constrained cost function must be optimized:

$$\arg \min_{\lambda} \{\mathbb{D}_{\text{KL}}(g, P\lambda) + \beta R(\lambda)\}, \quad (21)$$

where $\mathbb{D}_{\text{KL}}(g, P\lambda) = \sum_j (g_j \log \bar{g}_j + \bar{g}_j)$ is a Kullback–Leibler (KL) distance [5] and the second term is a convex energy functional controlled by the regularization parameter β .

In this work we consider two convex functionals for $R(\lambda)$. The first one is a discretization of the continuous functional (1):

$$R_1(\lambda) = \sum_i^N \phi(|\nabla \lambda_i|); \quad \phi(s) = -\frac{\epsilon^2}{2} \exp(-s^2/\epsilon^2). \quad (22)$$

The second penalty is given for a 2D image λ as:

$$R_2(\lambda) = \sum_i^N \sum_{k \in \mathbb{N}_i} \rho(\lambda_i - \lambda_k). \quad (23)$$

Penalty R_2 performs smoothing between the central pixel i and the nearest pixel k in the local neighbourhood set \mathbb{N}_i . By considering non-quadratic functions for ρ one can avoid boundaries blurring. Here we use a Huber function [19] with a threshold T :

$$\rho(x, T) = \begin{cases} x^2/2, & \text{if } |x| \leq T \\ T|x| - T^2/2 & \text{if } |x| > T. \end{cases} \quad (24)$$

It has been shown in [20] that the minimization of the functional (1) is equivalent to nonlinear tensor diffusivity PDE (12) with anisotropic tensor $\mathbf{D} = \frac{\phi'(|\nabla u|)}{|\nabla u|^3} \nabla u \nabla u^T$. The convexity of the functional R_1 is defined by the function ϕ which is differentiable and convex in our case (5). Since $\mathbb{D}_{\text{KL}}(g, P\lambda)$ is also a convex functional [5], one can consider the first order optimality condition for the problem (21). Using a Lagrange multiplier κ to ensure positivity constraint $\lambda > 0$ one can write the system:

$$\begin{cases} P^* \mathbf{1} - P^* \left(\frac{g}{P\lambda} \right) + \beta \partial_{\lambda} R\lambda - \kappa = 0 \\ \kappa \lambda = 0. \end{cases} \quad (25)$$

By multiplying (25) with λ one can obtain the following iteration scheme:

$$\lambda^{m+1} - \lambda^m \frac{\mathbf{P}^*}{\mathbf{P}^* \mathbf{1}} \left(\frac{\mathbf{g}}{\mathbf{P} \lambda^m} \right) + \beta \frac{\lambda^m}{\mathbf{P}^* \mathbf{1}} \partial_\lambda \mathbf{R} \lambda^m = 0. \quad (26)$$

The second term of (26) is a step of the MLEM method [4], therefore a nested iterative algorithm [13, 21] is given as:

$$\begin{aligned} \lambda^{m+\frac{1}{2}} &= \frac{\lambda^m}{\mathbf{P}^* \mathbf{1}} \mathbf{P}^* \left(\frac{\mathbf{g}}{\mathbf{P} \lambda^m} \right), & \text{MLEM step} \\ \lambda^{m+1} &= \mathbf{L}(\lambda^{m+\frac{1}{2}}), & \text{Denoising step.} \end{aligned} \quad (27)$$

Here MLEM method solves the KL optimization sub-problem and \mathbf{L} is an operator that performs a transition from $\lambda^{m+\frac{1}{2}}$ to λ^{m+1} by minimizing the following function:

$$\Psi(\lambda) = \frac{1}{2} \int_{\Omega} \frac{\mathbf{P}^* \mathbf{1}}{\lambda^m} (\lambda - \lambda^{m+\frac{1}{2}})^2 + \beta \mathbf{R}(\lambda). \quad (28)$$

The standard iterative gradient descent algorithm is used to optimize equation (28):

$$\begin{cases} \lambda^{l+1} = \lambda^l + \tau(\Psi'(\lambda)) \\ \lambda^{l=0} = \lambda^{m+\frac{1}{2}}. \end{cases} \quad (29)$$

Three different representations for the function $\Psi'(\lambda)$ were compared with each other in this work, namely:

$$\Psi'_1(\lambda) = \frac{\mathbf{P}^* \mathbf{1}}{\lambda^m} (\lambda - \lambda^{m+\frac{1}{2}}) + \beta (\nabla \cdot (\mathbf{D}_\lambda \nabla \lambda)) \quad (30)$$

$$\Psi'_2(\lambda) = \frac{\mathbf{P}^* \mathbf{1}}{\lambda^m} (\lambda - \lambda^{m+\frac{1}{2}}) + \beta (\nabla \cdot (\mathbf{D}_{\lambda, \mu} \nabla \lambda)) \quad (31)$$

$$\Psi'_3(\lambda) = \frac{\mathbf{P}^* \mathbf{1}}{\lambda^m} (\lambda - \lambda^{m+\frac{1}{2}}) + \beta \left(\sum_i^N \sum_{k \in \mathfrak{N}_i(\mu, n_0)} \rho'(\lambda_i - \lambda_k) \right). \quad (32)$$

The right hand side of the functional (30) includes the sum of the weighted data fidelity terms and the derivative of the penalty term \mathbf{R}_1 (22). The scalar term $\varphi(|\nabla \lambda|)$ (4) is replaced with the anisotropic diffusivity tensor (6). This is a weighted denoising step using the classical nonlinear anisotropic DT [9].

The functional (31) is identical to (30) except it implicitly uses the supplementary image μ in the proposed combined tensor framework (13). In the following experiments we used an identical regularization parameter β for (30) and (31) functionals, implying that the tensor framework remain the same only directions of smoothing can differ. Additionally by modifying (30) with the proposed combined tensor one can argue about the validity of minimized solution from the Bayesian perspective. However, this is the case with majority of methods using supplementary information [2]. The CDT is built using the image μ (e.g. representing an anatomical modality, such as MRI or CT) and emission image λ (see figure 1).

In this paper the penalty \mathbf{R}_2 (23) is coupled with a Bowsher method (BM) [22] which drives the smoothing process using the supplementary image μ . In the right hand side of (32) one can see the smoothing term which is conditioned on the neighbourhood $\mathfrak{N}_i(\mu, n_0)$. The neighbourhood depends on μ alone and n_0 is a number of the most closest neighbours of i based on the smallest absolute differences $|\mu_i - \mu_k|$. For more details on BM we refer to [2, 6].

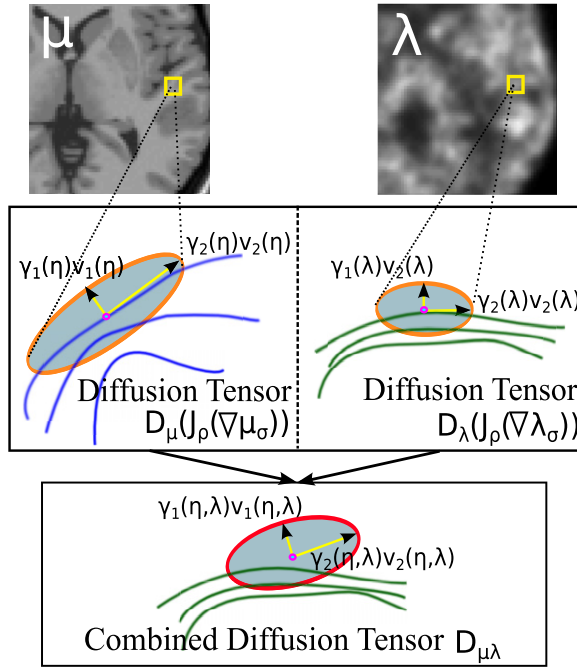


Figure 1. An example of local geometrical structure represented by two diffusion tensors for anatomical μ and emission images λ respectively. The combined diffusion tensor is constructed by interpolation.

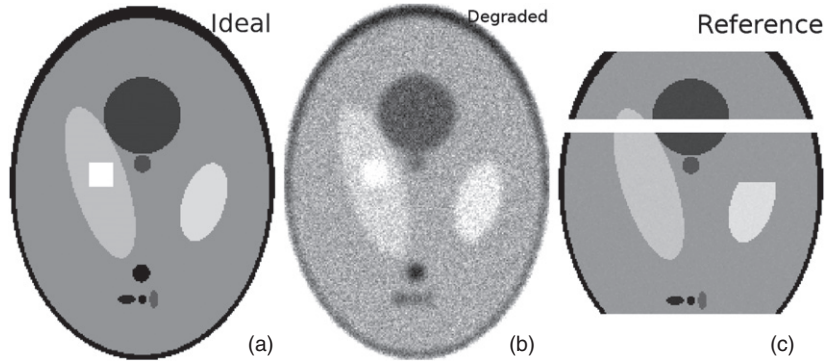


Figure 2. (a) Original image λ , (b) isotropically blurred and noisy image λ_{noise} with 12% of normally distributed random noise, (c) the reference image μ with 1% of noise. Some areas of μ are correlated with λ and some are intentionally geometrically different. The main goal of the recovery process of $\hat{\lambda}$ is to enhance mutual features existing in μ and λ without introducing the false edges caused by uncorrelated features.

3. Numerical results

3.1. Image restoration using CDT

Image λ (figure 2(a)) is the modified noiseless Shepp–Logan phantom and λ_{noise} is significantly degraded by an isotropic blur and 12% normally distributed random noise. Image μ (the reference) is a less noisy dataset (1% of the random noise) although with areas with

Table 1. Parameters for the image restoration experiment.

Method	σ_λ	σ_μ	ρ_λ	ρ_μ	ϵ_λ	ϵ_μ	ζ	T_λ	Δt	τ	l (19)	l (29)
DT	2.0	—	0.7	—	0.01	—	—	—	0.1	—	100	—
BM	—	—	—	—	—	—	—	0.01	—	0.1	—	20
CDT	2.0	2.0	0.7	0.6	0.01	0.04	0.0001	—	0.1	—	100	—

Table 2. NMSE values for DT, BM and CDT methods.

ROI	λ_{noise}	DT	BM	CDT
Whole phantom	0.1859	0.1614	0.157	0.159
Square and circle	0.2936	0.2015	0.2534	0.2012

no image information, making it geometrically different locally from λ , see figure 2(a). Image μ can also have different greyscale intensity values, although the two synthetic images used here have the same greyscale range to simplify the estimation of ϵ_μ , ϵ_λ and the comparison between methods. Our main aim here is to recover $\hat{\lambda}$ from λ_{noise} by using the available information contained in μ . Features of μ which are geometrically correlated (mutual edges) with λ must be enhanced in $\hat{\lambda}$, meanwhile absent or non correlated features of μ should be suppressed during the recovery process. Since some missing data in the reference image μ (e.g. the horizontal line across the image) can create sharp changes in intensity it is essential not to transfer this information into recovery of $\hat{\lambda}$. This is a challenging task and failing to do so will create severe artefacts in $\hat{\lambda}$. During recovery the major priority remains to preserve significant information in λ_{noise} .

For numerical experiments with DT and combined CDT we use iterative scheme (19). For the BM we use gradient descent iterations (29) (only upper equation) with the functional $\Psi'_3(\lambda)$ defined as $\Psi'_3(\lambda) = \sum_i^N \sum_{k \in \mathbb{N}_i(\mu, n_0)} \rho'(\lambda_i - \lambda_k)$. These methods require a few parameters which are provided in table 1.

Knowing the exact image one can find optimal parameters with a multidimensional optimization, which is a computationally expensive task. In this experiment all parameters were chosen empirically based on the known level of noise and the response from normalized mean square error (NMSE), given as:

$$\text{NMSE}(\hat{\lambda}, \lambda) = \frac{\|\hat{\lambda} - \lambda\|_2^2}{\|\lambda\|_2^2}. \quad (33)$$

The improvement in quality of the recovered image using supplementary information can be clearly seen for both BM (32) and CDT (31) methods (see figure 3(b) and (c) respectively)). We specify two different regions of interest (ROI) to calculate NMSE (33), namely, the whole phantom ROI and square and circle objects which are absent from μ (indicated by arrows in figure 3(b)). The values of (33) for the specified ROIs are provided in table 2.

The smallest NMSE value for the whole phantom ROI belongs to BM, it gives sharp edges and improved resolution in some parts of the phantom (see figure 3(b)). However the recovery overall is very noisy and some features are significantly distorted. The severity of artefacts is reflected in the higher values of NMSE for objects which do not exist in the supplementary image μ (indicated by arrows in figure 3(b)). The proposed CDT method gives a slightly higher value of NMSE (for the whole phantom) due to the blurring effect, however the recovery is almost noiseless and resolution is improved comparatively to DT method (see figure 3(c)). For the square and circle features which belong to λ only, both DT and CDT outperform BM in terms of NMSE (see table 2). The CDT method suggests a good compromise between the improved resolution and keeping important features in λ intact.

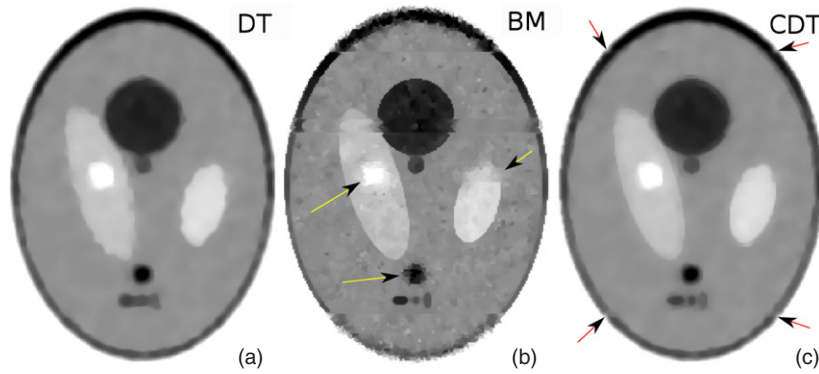


Figure 3. Recovered image $\hat{\lambda}$ using (a) DT filter (30), (b) BM (32) (c) CDT filter (31). In (b) the boundaries which are common to both λ and μ are enhanced significantly. However, in places where misalignment of data exist, very strong artefacts appeared and features which are absent in μ are corrupted (indicated by arrows). In (c) the mutual boundaries for both λ and μ are also enhanced. Some minor artifacts are visible on the outer shell of the phantom due to amplification of the false common edges (indicated by arrows).

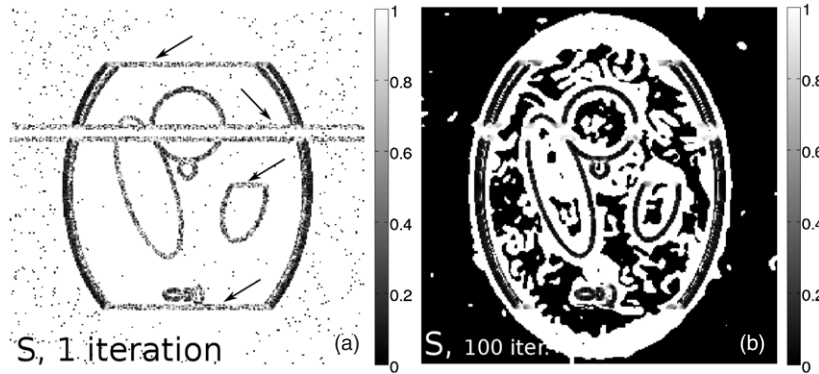


Figure 4. (a) Distribution of s after one iteration, showing the false common edges (indicated by arrows) determined due to noise. (b) distribution of s on the last iteration, showing the significant reduction of the false common edges due to performance of the successive steps 1 and 2.

After the first iteration of the proposed method (31), many false common edges are determined (see the s distribution in figure 4(a)), but on the last iteration the number of false common edges is reduced significantly (figure 4(b)). This behaviour is due to the step 2 where the gradient orientations for both images are estimated and compared [17]. In areas where no information exists in either images, the CDT acts isotropically. The s selection still identifies some false edges which introduce small artefacts in the image recovered with CDT (see arrows in figure 3(c)). Nevertheless, the resolution of features is improved and when true common edges are identified much sharper boundaries are reconstructed.

3.2. Numerical modelling of SPECT reconstruction using CDT

The idea of our CDT algorithm (13) is that it can be used for imaging modalities where one data set is supported by the co-registered set from a different modality. The data sets represent

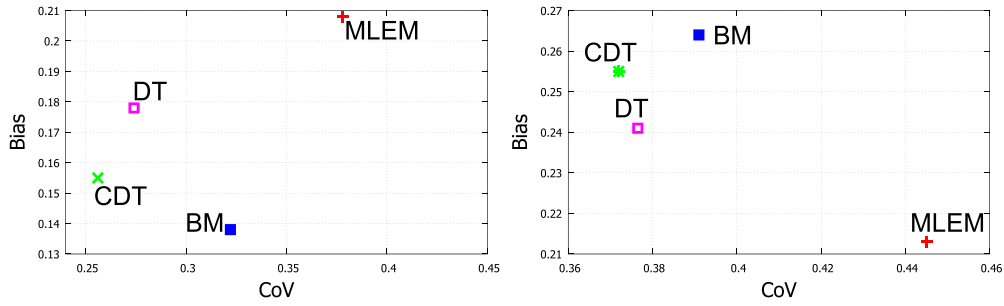


Figure 5. Bias-CoV (left) plots for the gray matter (GM) ROI and bias-CoV (right) for the hot lesion ROI (indicated by arrows in figure 6).

different physical properties, e.g. ET and x-ray CT or MRI where the emission data describe the functional behaviour of the biological tissue, and the CT/MRI show the anatomical structure at higher resolution [2].

For our experiments we used a 2D synthetic Brainweb dataset [23] to perform quantitative analysis of selected methods: MLEM (27) (upper step only), MLEM with BM (32), MLEM with DT (30), and MLEM with CDT (31). Since we perform the MLEM step similarly for every penalty function we reduce our notation for reconstruction methods with penalties to DT, BM, and CDT. The synthetic phantom (256×256 pixels) was projected with the SPECT system distance dependent point spread function to form a noise-free sinogram, then 20 Poisson noise realizations were generated. Several hot (HL) and cold lesions (CL) were added to the image phantom (see figure 6), but not to the reference MRI data, by increasing (HL) or decreasing (CL) the tracer uptake by 20% in the lesions. Systematic blur was added to the projections to model a typical SPECT distance dependent point spread function. The number of acquisition angles is 180 with 362 detectors.

For quantitative validation we calculate the bias-coefficient of variation (CoV) values in the hot lesion ROI (cold lesions were approximately the same in analysis) and the gray matter (GM). The bias between true activity λ and the estimated $\hat{\lambda}$, over a ROI, is defined as:

$$\text{bias}(\hat{\lambda}, \lambda) = \frac{1}{W} \sum_{w=1}^W \left(\frac{\hat{\lambda}_w - \bar{\lambda}}{\bar{\lambda}} \right)^2, \quad (34)$$

where W is a number of noise realizations and $\bar{\lambda}$ is a mean over all pixels in the ROI. The CoV is defined as:

$$\text{CoV}(\hat{\lambda}) = \frac{1}{|\text{ROI}|} \sum_{k \in \text{ROI}} \sqrt{\frac{\sum_{w=1}^W (\hat{\lambda}_{w,k} - \bar{\lambda}_k)^2}{(W-1)\bar{\lambda}_k^2}}, \quad (35)$$

where $\bar{\lambda}_{w,k}$ is the average of $\hat{\lambda}_k$ over W .

For the parameters given in table 3 we present the best achieved values of bias and CoV in figure 5. The values for the smoothing kernels σ_λ , σ_μ , ρ_λ , ρ_μ are taken to be the same as in table 1. To find an optimal value for regularization parameter β (21), given in table 3, we performed several tests with different values of β keeping the other parameters fixed (knowing the noise level we estimated them empirically). According to the bias-CoV response for different β we found approximate optimal values for each method.

Reconstruction of the activity phantom was performed using MLEM, DT, BM and CDT algorithm (see figure 6) using ideally co-registered MRI image. The MLEM image has a poor resolution and the highest CoV for both the GM and HL regions. The use of MLEM with DT

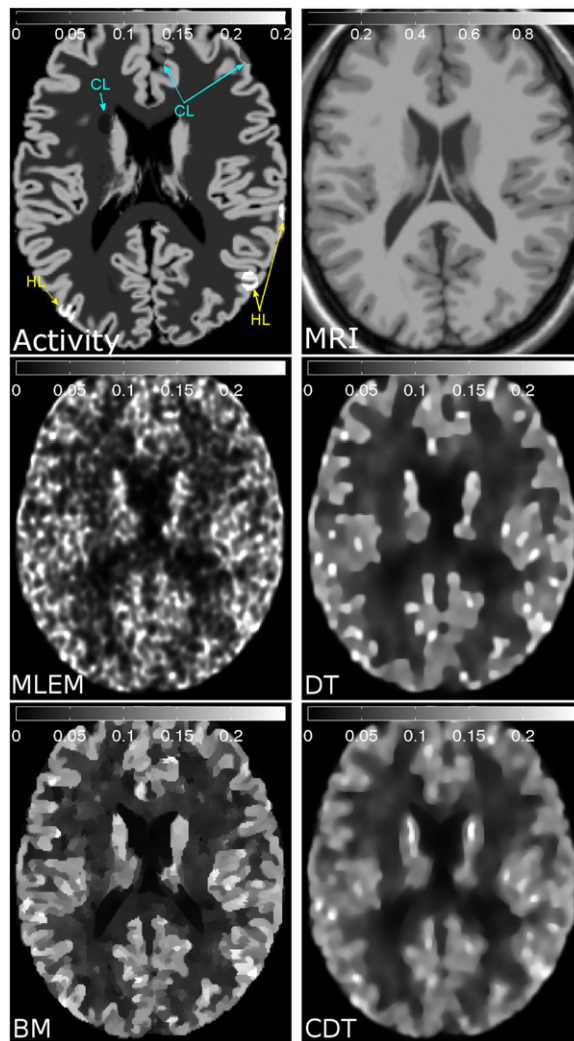


Figure 6. Left to right, top row: the 2D ideal and noiseless activity phantom with hot (HL) and cold lesions (CL) indicated by arrows, co-registered MR image; middle row: MLEM, DT; bottom row: BM, CDT.

Table 3. Parameters for the SPECT image reconstruction experiment.

Method	W	β	ϵ_λ	ϵ_μ	ζ	T_λ	Δt	τ	l (19)	m (27)
MLEM	20	—	—	—	—	—	—	—	—	110
DT	20	0.05	0.2	—	—	—	0.1	0.1	5	120
BM	20	0.3	—	—	—	0.01	—	0.1	2	120
CDT	20	0.05	0.2	0.04	0.0001	—	0.1	0.1	5	120

reconstruction gives much better noise removal than MLEM alone, but the bias level for GM is still high. The BM gives the lowest bias by strongly penalizing reconstruction of λ with information from μ , however the variance level is high due to noise. The CDT reconstruction

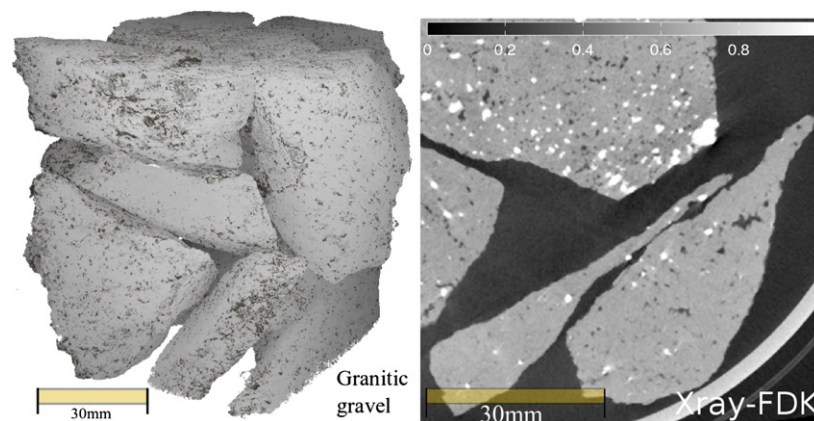


Figure 7. (Left) a 3D x-ray CT rendering of rock particles (an angular porous granitic gravel) inside a neutron transparent aluminium column. (Right) a 2D horizontal slice through the reconstructed x-ray volume, reconstruction was performed using the Feldkamp (FDK) method.

achieves better bias characteristics for GM comparatively to DT and reaches best variance level. The influence of anatomical data in CDT is clearly visible and results in a low bias for GM. For the lesion ROI the proposed method achieves also good performance without sacrificing bias significantly. Strong differences in intensity related to artificially imposed lesions experience minimal penalization since they contain valuable information. The overall quality of the image reconstructed with CDT is a good example of the trade-off between anatomical information and emission data. The BM also produces very good results, however it is slightly (high bias value for the lesion ROI) blurring the valuable information in λ and introducing more noise than the diffusion based methods. Potentially, the CDF method can be a good candidate for another anatomically driven smoothing technique for emission tomography [2].

3.3. Neutron data reconstruction using CDT

Using data obtained on the cold neutron imaging Beamline ICON at SINQ spallation neutron source at Paul Scherrer Institute, Switzerland [24] we can exploit the high neutron cross section of hydrogen to locate water rivulets within a gravel volume (see figure 7 (left)). Aqueous fluids are difficult to observe using x-rays because of the low x-ray cross section of hydrogen, similar to soft tissue in CT scans. The neutron imaging can give temporal information on water distribution within the rock, however the spatial resolution is significantly lower than the laboratory x-ray CT system (less than half). Exposure time is much longer, fluxes are generally lower, and NCT projections are more strongly influenced by a scattering contribution, making reconstructions significantly noisier than for x-ray CT. A quick scan of the sample with x-ray source before the neutron experiment can provide a dataset with superior spatial resolution. Using the technique presented one can combine neutron and x-ray tomographic modalities in iterative reconstruction of the neutron measurements.

In this experiment the laboratory x-ray tomography data was acquired on a Nikon XTH 225 ST at the Manchester X-ray Imaging Facility.

The NCT data formed part of a time lapse experiment which was primarily used to develop spatial-temporal regularization methods [25] to enable acceptable images to be obtained using increasingly short acquisition times. As such the NCT image acquisition is not optimized for quality and the images are noisy. In this experiment we used a 2D slice of x-ray dataset

Table 4. Parameters for the neutron data reconstruction experiment.

Method	β	ϵ_λ	ϵ_μ	ς	T_λ	Δt	τ	l (19)	m (27)
MLEM	—	—	—	—	—	—	—	—	100
DT	0.1	0.12	—	—	—	0.1	0.1	5	100
BM	0.55	—	—	—	0.2	—	0.1	3	100
CDT	0.1	0.12	0.04	0.0001	—	0.1	0.1	5	100

(see figure 7 (right)) as a structural prior for a ROI in the reconstruction of the NCT data. The x-ray data is cropped from the 3D rock sample (see figure 7 (left)) reconstructed from 2000 projections using a Feldkamp algorithm [26].

The set of 2D slices (1000×1000 pixels) using projection data obtained by NCT experiment were reconstructed from just 200 projections with MLEM, DT, BM and CDT algorithms (figure 8). For this experiment we used parameters given in table 4.

Since in this experiment we were interested in the dynamics of water flow through gravel, we demonstrate three different stages (one dry and two wet) to illustrate the reconstruction outputs. The spatial resolution of images reconstructed with MLEM is poor and images are noisy. The influence of noise is reduced by using the MLEM with DT, but the results have a blotchy appearance and low resolution. Incorporation of x-ray CT information with BM enhances the resolution of neutron reconstructions, however the noise level is high. The BM fails to smooth noise successfully due to local behaviour of the model. Moreover, if the reference image is noisy it is expected that BM will perform poorly (the absolute differences measure is very sensitive to noise). The CDT method significantly enhances rock interfaces without sacrificing features which are only observed in the neutron data. The method also reduces blotchiness which is visible on DT images. In this experiment an improved signal-to-noise ratio and resolution makes the CDT method the preferred option for accurate quantitative analysis.

4. Discussion

For the comparisons shown here we have minimized the number of reconstruction methods presented to allow us to focus on the novel incorporation of the prior. Further comparison to other structurally anatomically driven techniques, computational acceleration and 3D generalization are current research aims. We also limited the scope of this work to a method that performs Euclidean tensor interpolation [15] although other forms of interpolations, such as feature-based schemes based on Euler angles or quaternions [16] could also be employed.

Although the results using experimental data are encouraging, it should be noted that the registration accuracy is essential to avoid strong discrepancies. To minimize registration errors between data sets from different modalities it is advisable to perform scans in one gantry, such as hybrid medical scanners. When this is not possible, one can perform scans close to each other in time, this can reduce the possibility of the mismatch in the inner structure of the object or subject.

The performance of the CDT method is strongly dependant on the choice of s in the interpolation procedure. Steps 1 and 2 can be replaced with methods that are more robust to noise and blur when identifying joint edges and gradient orientations [18]. Most notably, the step involving gradient orientations comparison is crucial if artifacts in the reconstructed images are to be avoided. We have focussed on edge enhancing diffusion (11), but a coherence enhanced diffusion approach could also be implemented, or different eigenvalues for the DT

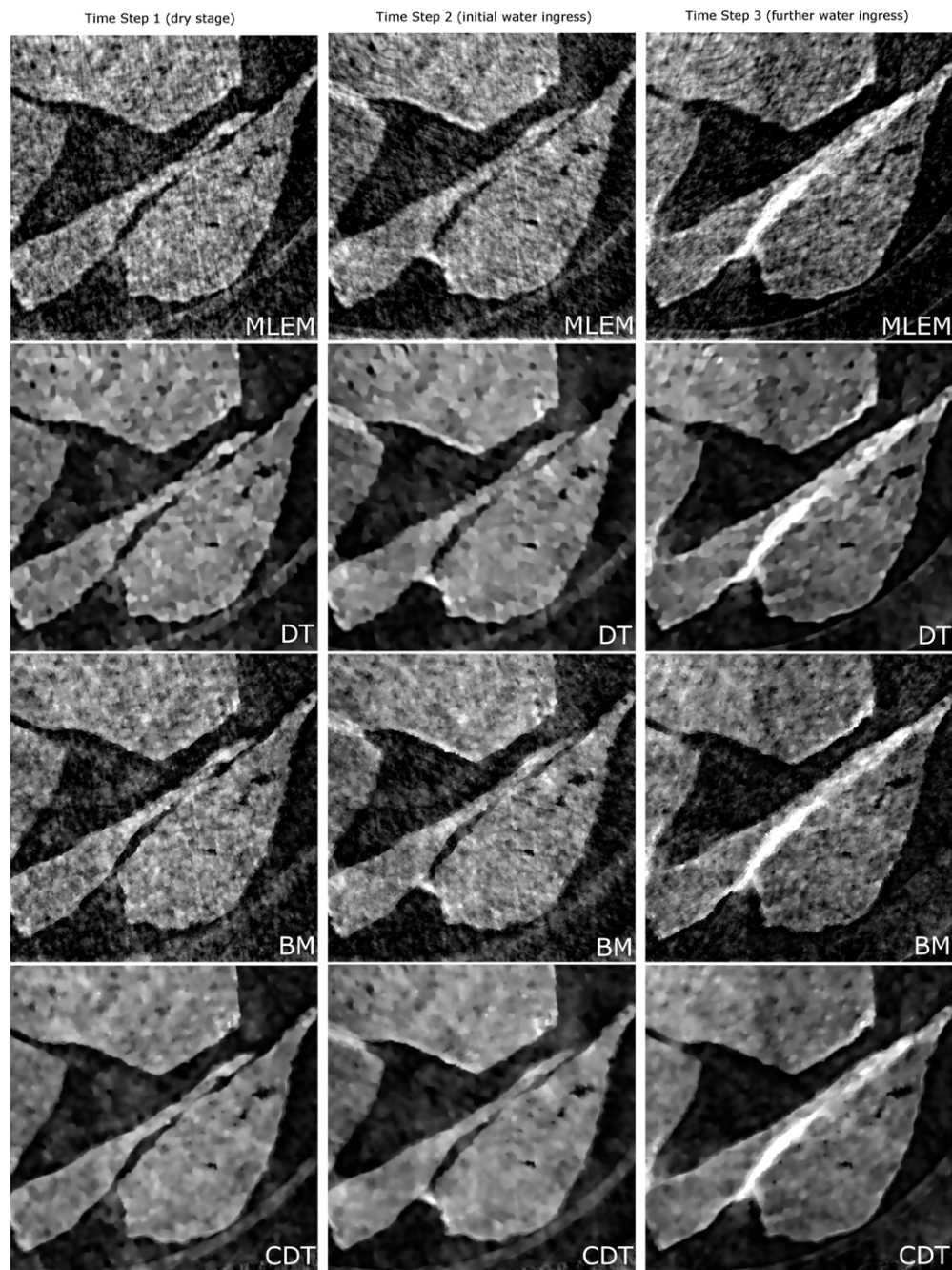


Figure 8. 2D slices of the porous granitic gravel sample showing the ‘dry’ stage (without water) and two subsequent ‘wet’ frames (with water distribution); top to bottom: reconstruction with MLEM method, DT (the noise level is significantly reduced), BM using the x-ray CT data as a prior (though the resolution is improved the noise level is high) and CDT method where resolution is improved and blotchiness is alleviated.

Table 5. Method performance comparison using Xeon CPU 2.26 GHz per one iteration l in (29).

Image size	BM	DT	CDT
1000^2	0.3 s	0.82 s	0.86 s

applied [9]. For example, the tangential eigenvalue (preferable direction of smoothing) can be dependent on the gradient's magnitude, and incorporating this into the treatment would make smoothing more sensitive to edges in this direction, but would reduce the spatial resolution of the reconstructed image.

One benefit of the approach shown here is that the combined tensor is a flexible and easy-to-use technique which produces satisfying results (both visually and quantitatively). The CDT can be constructed by interpolation from many given tensors (not just two) and this can be beneficial in temporal averaging problems or for using two images to provide structural information.

In terms of computational speed, the proposed method has almost the same performance as the conventional tensor based diffusivity filter (see table 5). Both DT and CDT methods have a good potential for parallelization on CPUs or GPUs.

5. Conclusions

We have shown a novel approach for incorporating available additional information from different imaging modalities into diffusion filtering processes. This information is embedded by scaling and rotating the combined diffusion tensor of two images. The method has the ability to detect matching directions of tensor fields in multiple images, allowing better edge identification and improved image resolution in the reconstructed image. The proposed method can be beneficial for quantification tasks where precise segmentation of certain classes is needed.

The proposed method was successfully applied to the 2D model case of a SPECT functional reconstruction with co-registered structural MRI data as the reference image. Information on functionally important regions like lesions is preserved while the anatomical boundaries are introduced. Reconstructed images look more favourable for subsequent image analysis due to a smoother appearance and good feature preservation.

The method was also applied to real noisy neutron tomography data. The neutron images are highly sensitive to the presence of water, but noisy and of poor resolution. By incorporating information from higher spatial resolution x-ray CT data it was possible to identify the regions of water clearly as well as obtain better information regarding the morphology of the gravel and pore spaces. This would enable significantly better quantification of the water ingress process.

Acknowledgments

This work has been supported by the Engineering and Physical Sciences Research Council under Grants EP/G026483/1, EP/J010456/1, EP/I02249X/1 and the CRUK/EPSCRC funded King's College/UCL Comprehensive Cancer Imaging Centre. UCL/UCLH receives a proportion of funding from the UK Department of Health's NIHR Biomedical Research Centre funding scheme. The authors thank the Research Complex at Harwell and the Paul Scherrer Institute for the use of their facilities.

References

- [1] Innis R B, Cunningham V J, Delforge J, Fujita M, Gjedde A, Gunn R N, Holden J, Houle S and Huang S C 2007 Consensus nomenclature for *in vivo* imaging of reversibly binding radioligands *J. Cereb. Blood Flow Metab.* **27** 1533–9
- [2] Vunckx K, Atre A, Baete K, Reilhac A, Deroose C M, Van Laere K and Nuyts J 2012 Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging *IEEE Trans. Med. Imaging* **31** 599–612
- [3] Hutton B F, Nuyts J and Zaidi H 2006 *Iterative Reconstruction Methods. Quantitative Analysis in Nuclear Medicine Imaging* (New York: Springer)
- [4] Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **MI-2** 113–22
- [5] Qi J and Leahy R M 2006 Iterative reconstruction techniques in emission computed tomography *Phys. Med. Biol.* **51** 541–78
- [6] Kazantsev D, Arridge S R, Pedemonte S, Bousse A, Erlandsson K, Hutton B F and Ourselin S 2012 An anatomically driven anisotropic diffusion filtering method for 3D SPECT reconstruction *Phys. Med. Biol.* **57** 3793–810
- [7] Chan C, Fulton R R, Feng D D and Meikle S R 2009 Regularized image reconstruction with an anatomically adaptive prior for positron emission tomography *Phys. Med. Biol.* **54** 7379–400
- [8] Perona P and Malik J 1990 Scale-space and edge detection using anisotropic diffusion *IEEE Trans. Pattern Anal.* **12** 629–39
- [9] Weickert J 1998 *Anisotropic Diffusion in Image Processing* (Stuttgart: Teubner)
- [10] Aubert G and Kornprobst P 2006 *Mathematical Problems in Image Processing. Partial Differential Equations and the Calculus of Variations* (New York: Springer)
- [11] Jahne B 1993 *Spatio-Temporal Image Processing: Theory and Scientific Applications* (Berlin: Springer)
- [12] Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica D* **60** 259–68
- [13] Sawatzky A, Brune C, Wubbeling F, Kosters T, Schafers K and Burger M 2008 Accurate EM-TV algorithm in PET with low SNR *IEEE Nuclear Science Symp. Conf. Record* pp 5133–7
- [14] Weickert J 1999 Coherence enhancing diffusion *Int. J. Comput. Vis.* **31** 111–27
- [15] Batchelor P G, Moakher M, Atkinson D, Calamante F and Connelly A 2005 A rigorous framework for diffusion tensor calculus *Magn. Reson. Med.* **53** 221–5
- [16] Yang F, Zhu Y M, Magnin I E, Luo J H, Croisille P and Kingsley P B 2012 Feature-based interpolation of diffusion tensor fields and application to human cardiac DT-MRI *Med. Image Anal.* **16** 459–81
- [17] Canny J 1986 A computational approach to edge detection *IEEE Trans. Pattern Anal. Mach. Intell.* **8** 679–98
- [18] Michelet F, Da Costa J P, Laviolle O, Berthoumieu Y, Baylou P and Germain C 2007 Estimating local multiple orientations *Signal Process.* **87** 1655–69
- [19] Huber P J 1981 *Robust Statistics* (New York: Wiley)
- [20] Tschumperle D 2002 PDEs based regularization of multivalued images and applications *PhD Thesis* University de Nice-Sophia Antipolis
- [21] Yan M, Bui A, Cong J and Vese L A 2013 General convergent expectation maximization (EM)-type algorithms for image reconstruction *Inverse Problems Imaging* **7** 1007–29
- [22] Bowsher J E et al 2004 Utilizing MRI information to estimate F18-FDG distributions in rat flank tumours *Proc. IEEE Nuclear Science Symp. Conf. Record* pp 2488–92
- [23] 2011 BrainWeb: Simulated brain database *Magn. Reson. Med.* BrainWeb <http://mouldy.bic.mni.mcgill.ca/brainweb/>
- [24] Kaestner A P, Hartmann S, Kuhne G, Frei G, Grunzweig C, Josic L, Schmid F and Lehmann E H 2011 The ICON beamline—a facility for cold neutron imaging at SINQ *Nucl. Instrum. Methods Phys. Res.* **659** 387–93
- [25] Kazantsev D, Lionheart W R B, Withers P J and Lee P D 2013 GPU accelerated 4D-CT reconstruction using higher order PDE regularization in spatial and temporal domains *CMSSE'13: Proc. Int. Conf. on Computational and Mathematical Methods in Science and Engineering (Almeria, Spain)* pp 843–52
- [26] Buzug T M 2008 *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT* (Berlin: Springer)